UNITED STATES
NAVAL POSTGRADUATE SCHOOL

THESIS

IDENTIFICATION OF LINEAR
SAMPLED DATA SYSTEMS

by

Ronald Keith Blackner

June 1967

IDENTIFICATION OF LINEAR

SAMPLED DATA SYSTEMS

by

Ronald Keith Blackner
Lieutenant Commander, United States Navy
B.S., Naval Academy, 1957

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ENGINEERING ELECTRONICS

from the

NAVAL POSTGRADUATE SCHOOL
June 1967

Signature of Author Ronald K Blackner

Approved by Harold Titus
Thesis Advisor

Charles H Rothauge
Chairman, Department of Electrical Engineering

R.F. Rinehart
Academic Dean

ABSTRACT

A least squares estimator $\hat{\Phi}$ is derived for the state transition matrix Φ of a linear, stationary sampled data system operating in a stochastic environment. The estimator $\hat{\Phi}$ is shown to be unbiased and minimum variance under the condition of full observability of the state vector of the system. The estimator is also shown to be the Maximum Likelihood Estimator for the case of the stochastic environment having Gaussian statistics. The estimation scheme is compared with two other recently published estimation schemes, both of which are shown to be special cases of the scheme herein presented.

TABLE OF CONTENTS

TABLE OF SYMBOLS

1. Vectors. (All lower case letters are vectors unless otherwise stated.)

$x(t)$	continuous representation of state vector
x_k	discrete representation of state vector
y_k	state vector of an equivalent system
z_k	observation of the system
w_k	system excitation
r_k	additive measurement noise
v_i	vector formed from a time sequence of a discrete state mode
ϵ_i	vector formed from a time sequence of an excitation mode
ϕ_i^T	i^{th} row vector in the matrix Φ

2. Matrices. (All upper case letters are matrices unless otherwise stated.)

F	matrix describing the homogeneous system
D	input matrix of the system
$\Phi(t)$	homogeneous solution of the system
Φ	state transition matrix of the discrete representation
Γ	input distribution matrix of the discrete system
J	matrix containing the cost function
X	matrix formed from a collection of state vectors
$\hat{\Phi}$	estimate of the state transition matrix
R	finite approximation to the autocorrelation function
P_m	inverse of R_o
H	observability transformation
Φ^*	state transition matrix of an equivalent linear system
Q	covariance of the random excitation

3. Scalar quantities.

L	quadratic cost function
$L(\phi_i, \sigma_i^2)$	log of the likelihood function of a random sample

n	system order
m	sample size
k	time index
α_m	a scalar function defined in the recursive form of the identification algorithm

4. Operators.

x^T	denotes x transpose
$(.)^{-1}$	denotes (.) inverse
$\nabla_x(.)$	denotes gradient of (.) with respect to x
$tr(.)$	denotes the trace of (.)
$E(.)$	denotes the expectation or mean of (.)
$cov(.)$	denotes the covariance of (.)
(a_{ij})	denotes A

1. INTRODUCTION

The problem to be considered is the identification, in a stochastic environment, of linear, stationary (constant coefficient), sampled data process dynamics. The problem is more concisely defined as follows:

Given:

1) A system whose behavior may be described by a set of linear, constant coefficient, differential equations.

2) A statistical description of the excitation of the system (the vector w in the figure below).

3) A sequence of observations on the state vector (x_k in the figure).

Problem:

Determine a "best" estimate of the system dynamics.

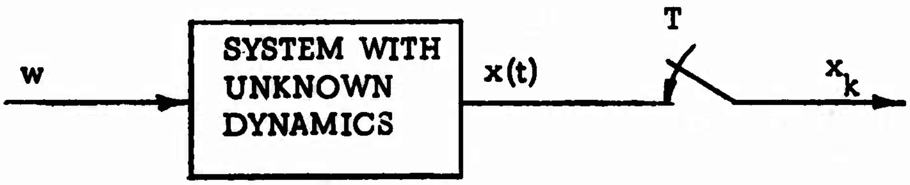


Figure 1. Schematic representation of the problem

Identification, defined in this manner, amounts to estimating the solution of the differential equation postulated in 1) above.

This problem may appear at the outset to be a rather restricted one. It becomes apparent, however, that a study of the problem takes on added significance when viewed in the context of a more general question. Given a vector of random time functions produced by some unknown process, can the process be adequately represented by a linear, stationary model, and if so, which linear model is the best representation of the system? This work is motivated primarily with the hope of finding application in the investigation of such questions.

Significant contributions in this area have been made by Ho and Whalen [2] and Lee [3]. The purpose of this work will be to extend

the methods suggested by the above authors. The approach here will be to derive a scheme which minimizes a quadratic cost criteria, subject to some rather limiting assumptions. An attempt will be made to analyze the consequence of relaxing these restrictive assumptions. Finally, the methods of Ho and Whalen [2] and Lee [3] will be generated from the method herein derived. Acknowledgement is given to LT Ralph E. Hudson, USN, who first suggested the algorithm to be presented.

2. BASIC MATHEMATICAL MODEL

Consider a dynamic system whose behavior may be defined by a set of n linear, constant coefficient, differential equations.

$$\dot{x} = F x + D w^*$$

The solution is given by

$$x(t) = \Phi(t-t_o)\, x_o + \int_{t_o}^{t} \Phi(t-\tau)\, D w^*(\tau)\, d\tau$$

where $\Phi(t-t_o)$ satisfies

$$\frac{d}{dt}\Phi(t-t_o) = F\,\Phi(t-t_o)$$

Introducing a sampling device of period T and a zero order hold on the excitation signal w*(t) makes possible the representation of the system at multiples of the sampling instant T by --

$$x_{k+1} = \Phi x_k + \Gamma w_k^*$$

where

$$x_k = x(kT)$$

$$\Phi = \Phi(T)$$

$$\Gamma = \Gamma(T) = \int_{0}^{T} \Phi(T-\tau) D\, d\tau$$

The process may be represented in block diagram as shown in Figure 2.

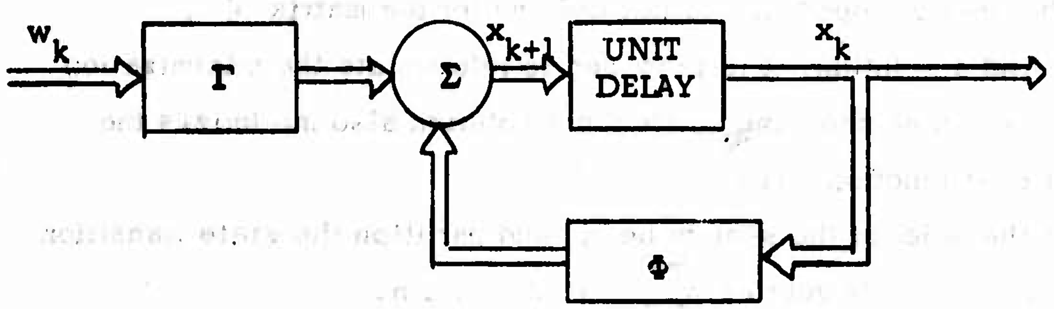


Figure 2. Schematic representation of the difference equation

3. DERIVATION OF THE IDENTIFICATION ALGORITHM

Consider a process of the form

$$x_{k+1} = \Phi x_k + w_k \tag{3.1}$$

$$w_k = \Gamma w_k^*$$

Here w_k is defined to be a gaussian sequence of zero mean, with covariance matrix Q. Suppose that a sequence of m+1 observations are made on the state vector x; $(x_o, x_1, \ldots x_m)$, from which it is desired to form an estimate of the state transition matrix Φ, assumed to be unknown. The so-called "least squares" estimator is that estimator $\hat{\Phi}$ which minimizes the quadratic cost function L (a scalar), where

$$L = \sum_{k=0}^{n-1} (x_{k+1} - \hat{\Phi} x_k)^T (x_{k+1} - \hat{\Phi} x_k) \tag{3.2}$$

Equivalently, $\hat{\Phi}$ minimizes the trace of J, where

$$J = \sum_{k=0}^{n-1} (x_{k+1} - \hat{\Phi} x_k)(x_{k+1} - \hat{\Phi} x_k)^T \tag{3.3}$$

From inspection, $L = tr(J)$.

The classical method of solving a problem of this nature is to set the gradient of the cost function with respect to the estimator $\hat{\Phi}$ equal

to zero. The solution to this gradient equation is sufficient to establish a minimum cost. The method cannot be applied here, however, since the gradient operation is not defined for the matrix $\hat{\Phi}$.

To find a solution, it is proposed to reformulate the minimization criteria, and then show the subsequent solution also minimizes the original cost function, tr(J).

Let the order of the system be n, and partition the state transition matrix Φ into n row vectors ϕ_i^T , i=1, 2, . . . n.

Then the process model

$$x_{k+1} = \Phi x_k + w_k \tag{3.4}$$

may be written as m scalar equations,

$$x_{k+1}^i = \phi_i^T x_k + w_k^i \ . \tag{3.5}$$

Here the superscript denotes the i^{th} scalar component (or mode) of the vector. From the m+1 vector observations, $(x_o, x_1, \ldots x_m)$, define

$$v_i = \begin{bmatrix} x_1^i \\ x_2^i \\ \cdot \\ \cdot \\ \cdot \\ x_n^i \end{bmatrix} \qquad \epsilon_i = \begin{bmatrix} w_0^i \\ w_1^i \\ \cdot \\ \cdot \\ \cdot \\ w_{n-1}^i \end{bmatrix} \qquad X^T = [x_0 x_1 \ldots x_{m-1}]$$

(mx1 vector) (mx1 vector) (nxm matrix)

From (3.5), these newly defined terms are related by

$$v_i = X\phi_i + \epsilon_i \tag{3.6}$$

Now consider an estimator $\hat{\phi}_i$, which minimizes the quadratic cost function g_i (a scalar), where

$$g_i = (v_i - X\hat{\phi}_i)^T (v_i - X\hat{\phi}_i) \tag{3.7}$$

Comparing equations 3.3 and 3.7, it is seen that

$$g_i = j_{ii}\text{, where } J = (j_{ik})$$

Thus, the sum of the new cost function g_i, i=1, 2, . . . n, is precisely the same as that previously defined, in that

$$\sum_{i=1}^{n} g_i = \mathrm{tr}(J) \tag{3.8}$$

A sufficient condition for the minimization of tr(J), then, is that g_i be minimized for i=1, 2, . . . n. Further, g_i is not a function of $\hat{\phi}_j$, $i \neq j$. This implies that the optimum estimator $\hat{\phi}_i$ does not depend upon the choice of the estimator selected for the jth row of the matrix estimator, $\hat{\Phi}$. Thus, each $\hat{\phi}_i$ may be found independently.

To accomplish the minimization, expand equation 3.7 and perform the gradient operation.

$$\nabla_{\hat{\phi}_i} g_i = -(X^T v_i)^T - v_i^T X + 2\,\hat{\phi}_i^T X^T X \tag{3.8}$$

Setting $\nabla_{\hat{\phi}_i} g_i = 0$ and solving for $\hat{\phi}_i$

$$\hat{\phi}_i^T = v_i^T X(X^T X)^{-1} \tag{3.9}$$

Equation (3.9) assumes $X^T X$ to be nonsingular. The conditions for nonsingularity are discussed in Section 5.

The optimum estimator $\hat{\Phi}$ can now be formed by placing the row vectors $\hat{\phi}_i^T$ in matrix form.

$$\hat{\Phi} = \begin{bmatrix} \hat{\phi}_1^T \\ \hat{\phi}_2^T \\ \vdots \\ \hat{\phi}_m^T \end{bmatrix} = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_m^T \end{bmatrix} X(X^1X)^{-1} \tag{3.10}$$

Finally, writing $\hat{\Phi}$ as a function of the original sequence of vector observations $x_o, x_1, \ldots x_m$, gives

$$\hat{\Phi} = \sum_{k=0}^{m-1} x_{k+1} x_k^T \left[\sum_{k=0}^{m-1} x_k x_k^T \right]^{-1} \tag{3.11}$$

Here $\sum_{k=0}^{m-1} x_k x_k^T$ is recognized as a finite approximation to the autocorrelation function, R(kT), evaluated at k=0. Similarly, $\sum_{k=0}^{m-1} x_{k+1} x_k^T$ is a finite approximation to R(T). Making this identification

$$\hat{\Phi} = R_1 (R_o)^{-1} \tag{3.12}$$

where

$$R_1 = R(T) \qquad\qquad R_o = R(0) \quad .$$

4. DEVELOPMENT OF A RECURSIVE FORM

Note that (3.12) gives the optimum estimator $\hat{\Phi}$, based on a sample of size m+1. The estimator is, then, a function of m, allowing the functional notation

$$\hat{\Phi}_m = R_1(m) [R_o(m)]^{-1} \tag{4.1}$$

To keep the index consistent with the sample size, define

$$R_1(m) \equiv \sum_{k=0}^{m-1} x_{k+1} x_k^T$$

$$R_o(m) \equiv \sum_{k=0}^{m-1} x_k x_k^T$$

Incrementing the index in (4.1) gives

$$\hat{\Phi}_{m+1} = [R_1(m+1)][R_o(m+1)]^{-1}$$

From the definition of $R_1(m)$

$$\hat{\Phi}_{m+1} = [R_1(m) + x_{m+1} x_m^T][R_o(m) + x_m x_m^T]^{-1} \qquad (4.2)$$

To ease the reader through the remainder of the recursive development, let

$$x \equiv x_m \qquad x_1 \equiv x_{m+1} \qquad R_o \equiv R_o(m)$$

Then

$$\hat{\Phi}_{m+1} = (R_1 + x_1 x^T)(R_o + x x^T)^{-1}$$

Appealing to the matrix inversion lemma, [5]

$$(R_o + x x^T)^{-1} = R_o^{-1}(I - x[x^T R_o^{-1} x + 1]^{-1} x^T R_o^{-1}) \qquad (4.3)$$

Recognizing $R_1 R_o^{-1}$ as $\hat{\Phi}_m$

$$\hat{\Phi}_{m+1} = [\hat{\Phi}_m + x_1 x^T R_o^{-1}][I - x(x^T R_o^{-1} x + 1)^{-1} x^T R_o^{-1}]$$

Define the scalar $\alpha_m = (x_m' R_o^{-1} x_m + 1)^{-1}$

Collecting terms and simplifying

$$\hat{\Phi}_{m+1} = \hat{\Phi}_m + \alpha_m (x_{m+1} - \hat{\Phi}_m x_m) x_m^T R_o^{-1}(m) \qquad (4.4)$$

Define $P_m = R_o^{-1}(m)$.

From equations (4.3) and (4.4), the final recursive forms can be written down.

$$\hat{\Phi}_{m+1} = \hat{\Phi}_m + \alpha_m (x_{m+1} - \hat{\Phi}_m x_m) x_m^T P_m \qquad (4.5)$$

$$P_{m+1} = P_m(I - \alpha_m x_m x_m^T P_m) \qquad (4.6)$$

$$\alpha_m = (x_m P_m x_m^T + 1)^{-1} \qquad (4.7)$$

Implicit in the use of these recursive forms is the utilization of (4.1) to generate an initial estimator $\hat{\Phi}$ and an initial matrix P_m. This fact would seriously complicate a mechanization of the iterative process. However, this difficulty can be avoided by the appropriate assignment of values to the initial estimator $\hat{\Phi}$, and the initial P_m matrix. Digital simulation has produced no apparent degradation of the estimation scheme when using this ruse. This subject is further discussed in Section 10.

5. NECESSARY AND SUFFICIENT CONDITIONS FOR IDENTIFIABILITY

Equation (3.9) assumed the nonsingularity of

$$R_o = \sum_{k=0}^{m-1} x_k x_k^T \qquad \text{(nxn) matrix}$$

Consider the decomposition of $R_o = X^T X$

$$X^T = \begin{bmatrix} x_o^1 & x_1^1 & x_2^1 & \cdots & x_{m-1}^1 \\ x_o^2 & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ x_o^n & & & & x_{m-1}^n \end{bmatrix} \qquad \text{(nxm) matrix}$$

Here the superscript again denotes the element or mode of the state vector.

From the fact that X must be of rank n if R_o is to be non-singular, it is immediately apparent that $m \geq n$ is a necessary condition for

identifiability. The number of observations on the state vector x_k must be equal to or greater than the system order, n. From inspection of X, it is seen that if any mode of the system, x_k^i, $k = 0, 1, \dots m-1$, remains at zero for all k, or if any two modes remain at some constant value, X will be of rank $r < n$, and R_o will be singular. These conditions imply that all modes of the system must be excited if R_o^{-1} is to exist.

Finally, to insure sufficiency of the above conditions, recall that the solution of an n^{th} order differential equation generates exactly n linearly independent solutions. The solution for the n modes of the system, then, form sets of linearly independent vectors, provided all modes are present.

From the fact that R_o has the decomposition X^TX, it follows that it is positive semidefinite. If the system is identifiable, R_o will be nonsingular and therefore positive definite. R_o is further seen to be symmetric, from the decomposition X^TX.

6. IDENTIFICATION OF A SYSTEM WITH CONSTRAINED OBSERVABILITY

The material presented in this section is an extension of a development of Lee [3], who considers the case of a scalar observable. Consider first a free (unforced) linear process of order n.

$$x_{k+1} = \Phi x_k = \Phi^{k+1} x_o \tag{6.1}$$

Let the observability of the state vector be constrained

$$z_k = H x_k \tag{6.2}$$

Here z_k, the observation, is an $(\ell \times 1)$ vector, $\ell < n$, and H, the observability transformation is of dimension $(\ell \times n)$. Such a system is said to be observable if the initial state vector x_o may be determined from the sequence of observations $z_o, z_1, \dots$ assuming that H and Φ are known. Forming this sequence of observations into an (nx1) vector gives

$$y_o = \begin{bmatrix} z_o \\ z_1 \\ \cdot \\ \cdot \\ z_\nu \end{bmatrix} \tag{6.3}$$

where z_ν denotes that observation required to make y_o of dimension (nx1). For example, if n=5 and ℓ=2

$$y_o = \begin{bmatrix} z_o^1 \\ z_o^2 \\ z_1^1 \\ z_1^2 \\ z_2^1 \end{bmatrix}$$

Here z_ν denotes the element z_2^1.

From (6.1) and (6.2)

$$y_o = \begin{bmatrix} z_o \\ z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_\nu \end{bmatrix} = \begin{bmatrix} H x_o \\ H \Phi x_o \\ \cdot \\ \cdot \\ \cdot \\ H_\nu \Phi^\nu x_o \end{bmatrix} = \begin{bmatrix} H \\ H \Phi \\ \cdot \\ \cdot \\ \cdot \\ H_\nu \Phi^\nu \end{bmatrix} x_o = A x_o \tag{6.4}$$

$H_\nu \Phi^\nu$ is defined in the same context as z_ν above, and may be written out for the example of m=5, ℓ=2.

$$H \Phi^2 = \begin{bmatrix} H_\nu \Phi^\nu \\ \text{------} \\ B \end{bmatrix}$$

Referring to (6.4), it is seen that the unique determination of the vector x_o from the sequence of observations $z_o, z_1, \ldots z_\nu$, is dependent upon the nonsingularity of the matrix A. If A is nonsingular, H, Φ are said to constitute an observable pair.

Assume H, Φ are an observable pair, and consider the forced system

$$x_{k+1} = \Phi x_k + \Gamma w_k^* \tag{6.5}$$

Define a new vector

$$y_k = A x_k \tag{6.6}$$

where the transformation A is defined by (6.4).

Incrementing the index gives

$$y_{k+1} = A x_{k+1} = A (\Phi x_k + \Gamma w_k^*) \tag{6.7}$$

Since A is by assumption nonsingular

$$y_{k+1} = A \Phi A^{-1} y_k + A \Gamma w_k^* \tag{6.8}$$

Define $\quad \Phi^* = A \Phi A^{-1}$ (6.9)

From the definition of (6.9), Φ and Φ^* are similar matrices, and will therefore have identical eigenvalues. Further, if the process is constrained with an observability matrix H

$$z_k = H x_k = H A^{-1} y_k \equiv H^* y_k . \tag{6.10}$$

From (6.10), note the sequence of observations $z_o, z_1, \ldots$ will be identical if generated by either of the processes

$$\begin{aligned} y_{k+1} &= \Phi^* y_k + A \Gamma w_k^* \\ z_k &= H^* y_k \end{aligned} \tag{6.11}$$

or

$$\begin{aligned} x_{k+1} &= \Phi x_k + \Gamma w_k^* \\ z_k &= H x_k \end{aligned} \tag{6.12}$$

Φ^* and H^* can be shown to be of special form. To demonstrate, consider again a 5^{th} order system with two system modes observable. That is, z_k a (2x1) vector, and x_k a (5x1) vector. From (6.9),

$$\Phi^* = A\,\Phi\,A^{-1} = \begin{bmatrix} H \\ H\,\Phi \\ H_\nu\,\Phi^\nu \end{bmatrix} \Phi \begin{bmatrix} H \\ H\,\Phi \\ H_\nu\,\Phi^\nu \end{bmatrix}^{-1} \tag{6.13}$$

where, as before,

$$H\,\Phi^2 = \begin{bmatrix} H_\nu\,\Phi^\nu \\ \hline B \end{bmatrix}$$

Taking Φ into the left bracket of (6.13) gives

$$\Phi^* = \begin{bmatrix} H\,\Phi \\ H\,\Phi^2 \\ (H_\nu\,\Phi^\nu)\,\Phi \end{bmatrix} \begin{bmatrix} H \\ H\,\Phi \\ H_\nu\,\Phi^\nu \end{bmatrix}^{-1} = \begin{bmatrix} C \\ \hline E \end{bmatrix} \begin{bmatrix} H \\ \hline C \end{bmatrix}^{-1}$$

$$\Phi^* = \begin{bmatrix} C \\ \hline E \end{bmatrix} [H_{-1} \quad C_{-1}] \tag{6.14}$$

where

$$[H_{-1} \quad C_{-1}] = \begin{bmatrix} H \\ \hline C \end{bmatrix}^{-1} \tag{6.15}$$

From equation (6.15)

$$[H_{-1}\ H] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad [C_{-1}\ C] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$[H_{-1}\ C] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad [C_{-1}\ H] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

performing the multiplication in (6.14) gives

$$\Phi^* = \left[\begin{array}{cc|c} 0 & 0 & \\ 0 & 0 & I \\ 0 & 0 & \\ \hline E & H_{-1} & E\ C_{-1} \end{array}\right] \tag{6.16}$$

In general, by straightforward extension of the above procedure

$$\Phi^* = \left[\begin{array}{c|c} 0 \cdots & \\ \vdots & I \\ \hline \multicolumn{2}{c}{\phi^{*T}_{n-\ell+1}} \\ \multicolumn{2}{c}{\vdots} \\ \multicolumn{2}{c}{\phi^{*T}_{n}} \end{array}\right] \tag{6.17}$$

H* can be found using (6.10) and (6.15), for the example case of H a (2x5) matrix.

$$H^* = H A^{-1} = H\,[H_{-1}\ \ C_{-1}] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

For the general case

$$H^* = \left[\begin{array}{c|ccc} I & 0 & \cdot & \cdot & \cdot \\ & \cdot & & & \\ & \cdot & & & \end{array}\right] \tag{6.18}$$

where H^* is of dimension $(\ell \times n)$.

Reconsider equation (6.11) for the special case where the first $(n-\ell)$ elements of the vector $A\,\Gamma\,w_k^*$ are identically zero for all w_k^*.

$$y_{k+1} = \Phi^*\, y_k + A\,\Gamma\,w_k^*$$

$$z_k = H^*\, y_k$$

A moment's reflection of the special form of Φ^* and H^* will show

$$y_k = \begin{bmatrix} z_k \\ z_{k+1} \\ \cdot \\ \cdot \\ z_{k+\nu} \end{bmatrix} \tag{6.19}$$

Thus, for this special case (the conditions being on $A\,\Gamma$), the complete state vector of an equivalent linear system Φ^*, H^*, may be constructed from the observations z_k, provided only that H, Φ form an observable pair. The case of nonzero elements in the vector $A\,\Gamma\,w_k^*$ is discussed in Section 8.

7. SOME PROPERTIES OF THE ESTIMATOR

This section is concerned with some properties of the estimator, $\hat{\Phi}$. Recall (3.6) and (3.9),

$$v_i = X\,\phi_i + \epsilon_i \tag{3.6}$$

$$\hat{\phi}_i = v_i^T X (X^T X)^{-1} = v_i^T X R_o^{-1} \quad . \tag{3.9}$$

This formulation fits nicely into the framework of classical regression analysis [4]. This fact can be utilized to determine several

statistical properties of the estimator, $\hat{\phi}_i$. In particular, the expected value of the estimator may be calculated. Recall the vector ϵ_i was formed from a scalar sequence of zero mean.

$$\epsilon_i = \begin{bmatrix} w_o^i \\ \cdot \\ \cdot \\ w_n^i \end{bmatrix} \quad ; \qquad E(w_k^i) = 0$$

From (3.6) and (3.9)

$$E(v_i) = E(X\,\phi_i + \epsilon_i) = X\,\phi_i \tag{7.1}$$

$$E(\hat{\phi}_i^T) = E(v_i^T X (X'X)^{-1}) = \phi_i^T \tag{7.2}$$

If ϵ_i has non zero mean, say $E(\epsilon_i) = \delta_i$, then

$$E(\hat{\phi}_i) = \phi_i + \delta_i X R_o^{-1} \tag{7.3}$$

A requirement, then, that $\hat{\phi}_i$ be an unbiased estimator is that the excitation be of zero mean.

To calculate the covariance of the estimator

$$\text{cov}(\hat{\phi}_i) = E(\hat{\phi}_i - E\,\hat{\phi}_i)(\hat{\phi}_i - E\,\hat{\phi}_i)^T \tag{7.4}$$

note that

$$\hat{\phi}_i - E\,\hat{\phi}_i = (X^TX)^{-1}X^T(v_i - E\,v_i) \tag{7.5}$$

$$v_i - E\,v_i = \epsilon_i$$

$$\text{cov}(v_i) = E(\epsilon_i\,\epsilon_i^T) = \sigma_i^2\,I \quad (\sigma_i^2 \text{ a scalar}) \tag{7.6}$$

If w_k^i is formed from the sampling of a band limited spectrum, (7.6) will not hold, since it assumes

$$E(w_k^i w_j^i) = \begin{cases} 0 & j \neq k \\ \sigma_i^2 & j = k \end{cases}$$

σ_i^2 is further seen to be equal to q_{ii}, where $Q = \text{cov}(w_k)$. Assuming w_k^i to be from a white spectrum gives, using (7.5) and (7.6)

$$\text{cov}(\hat{\phi}_i) = (X R_o^{-1})^T E(\epsilon_i \epsilon_i^T) X R_o^{-1} = \sigma_i^2 R_o^{-1} \tag{7.7}$$

For the case of ϵ_i a gaussian sample of zero mean, one can write out the joint distribution (likelihood) function of the sample. Designate the log of this function by $L(\phi_i, \sigma_i^2)$.

$$L(\phi_i, \sigma_i^2) = -\tfrac{1}{2}m(\log 2 + \log \sigma_i^2)$$

$$- \tfrac{1}{2}(v_i - X\phi_i)^T(v_i - X\phi_i) / \sigma_i^2 \tag{7.8}$$

The maximum likelihood estimator is formed by setting

$$\nabla_{\phi_i} L(\phi_i, \sigma_i^2) = 0$$

Note however, from equation (3.7), that the gradient of the original cost function g_i, is precisely the gradient of $L(\phi_i, \sigma_i^2)$. Clearly, then, $\hat{\phi}_i$ is the maximum likelihood estimator in the case of gaussian, zero mean excitation.

To find the maximum likelihood estimator for the variance of the excitation, σ_i^2, set

$$\nabla_{\sigma_i^2} L(\phi_i, \sigma_i^2) = 0$$

$$\hat{\sigma}_i^2 = (v_i - X\phi_i)^T(v_i - X\phi_i)/m$$

Replacing ϕ_i with $\hat{\phi}_i$ in (7.8), and using (3.9), this expression reduces to

$$\hat{\sigma}_i^2 = (v_i - \hat{\phi}_i X)^T v_i/m \tag{7.9}$$

This same estimator can be derived from the sample variance of w_k, since

$$\hat{Q} = \sum_{k=0}^{m-1} (w_k w_k^T)/m = \sum_{k=0}^{m-1} (x_{k+1} - \Phi x_k)(x_{k+1} - \Phi x_k)^T/m \tag{7.10}$$

Replacing Φ with $\hat{\Phi}$, and using the fact that

$$\hat{\Phi} X^T X \hat{\Phi}^T = R_1 \hat{\Phi}^T$$

gives

$$\hat{Q} = \sum_{k=0}^{m-1} \frac{1}{m}(w_k w_k^T) \approx (R_o - \hat{\Phi} R_1^T)/m \tag{7.11}$$

This estimator must converge to Q, since $\hat{\Phi}$ is unbiased. Note that the same result for (q_{ii}) is given by (7.9). For the case of Γ an n vector and w_k^* a scalar, (7.11) can be used to estimate Γ, for in this case,

$$Q = E(\Gamma w_k^* w_k^T \Gamma^T) = \sigma^2 \Gamma\Gamma^T \tag{7.12}$$

The properties of the estimator $\hat{\Phi}$ are summarized as follows.

1) $\hat{\Phi}$ is the minimum variance, unbiased estimator of Φ, provided the mean of the excitation is zero.
2) For the case of gaussian, zero mean excitation, $\hat{\Phi}$ is the maximum likelihood estimator of Φ.
3) For the case of gaussian, zero mean excitation,

$$\text{cov}(\phi_i) = \sigma_i^2 R_o^{-1}$$

$$\hat{Q} \approx (R_o - \hat{\Phi} R_1^T)/m$$

8. IDENTIFICATION WITH NOISY AND CONSTRAINED OBSERVATIONS

Consider first a system whose complete state vector may be observed, but the observations are contaminated by additive noise. Let the observation noise be from a gaussian, zero mean distribution.

$$x_{k+1} = \Phi x_k + w_k$$

$$z_k = x_k + r_k \tag{8.1}$$

In consonance with the definitions of section 3, define

$$\rho_i = \begin{bmatrix} r_1^i \\ r_2^i \\ \cdot \\ \cdot \\ r_n^i \end{bmatrix} \qquad R^T = (r_o \; r_1 \; . \; . \; . \; r_{m-1})$$

(mx1 vector) (nxm matrix)

Defining v_i, X, and ϵ_i as in section 3 gives

$$v_i = X \phi_i + \epsilon_i \tag{3.6}$$

Introducing the new variables

$$u_i = v_i + \rho_i$$

$$W = X + R$$

gives, from equation (3.6)

$$u_i - \rho_i = (W - R) \phi_i + \epsilon_i \tag{8.2}$$

Note here that the arrays u_i and W may be formed from the sequence of noisy observations, z_k.

Rearranging (8.2) gives

$$u_i = W \phi_i + \epsilon_i + \rho_i - R \phi_i$$

Now define a new vector

$$\xi_i = \epsilon_i + \rho_i - R \phi_i$$

A typical element of this vector is given by

$$\xi_i^k = w_k^i + r_k^i - r_{k-1}^T \phi_i$$

From the assumed distribution of the random vectors r and w

$$E(r_k r_j^T) = 0 \qquad \text{for } j \neq k$$

$$E(w_k r_j^T) = 0 \qquad \text{for all } j, k$$

Thus ξ_i is the sum of independent, zero mean gaussian vectors, and is therefore gaussian with zero mean. The variance of the vector is the sum of the variances of its component parts,

$$E(\xi_i \xi_i^T) = \sigma_{\xi_i}^2 I = \left[\sigma_{w_i}^2 + \sigma_{r_i}^2 + \phi_i^T \phi_i \sum_{j=1}^{m} \sigma_{r_j}^2\right] I$$

and from (8.2)

$$u_i = W \phi_i^T + \xi_i \tag{8.3}$$

From the form of equation (8.3), and the conditions on ξ_i it is apparent that all the preceeding analysis leading to an optimum identification scheme is also valid for the case of noisy observations. The deleterious effect of the measurement noise is readily apparent in the increased covariance of the estimator.

$$\mathrm{cov}(\hat{\phi}_i) = \sigma_{\xi_i}^2 (W^T W)^{-1}$$

The formulation also applies to estimation with constrained observations, as discussed in section 6.

$$y_{k+1} = \Phi^* y_k + A \Gamma w_k$$

$$z_k = H^* y_k$$

Here z_k is an (ℓ x1) vector, $\ell < n$. Making the assumption that the first (n-ℓ) elements of the vector $A \Gamma w_k^*$ were identically zero, it was shown

$$y_k = \begin{bmatrix} z_k \\ z_{k+1} \\ \cdot \\ \cdot \\ z_{k+\nu} \end{bmatrix} \tag{6.19}$$

For the more general case where $A\,\Gamma\, w_k^*$ has non zero elements, consider again the example of a 5^{th} order system with 2 observables. Let Γ be a vector distribution matrix, and w_k^* a scalar. Let the elements of the vector $A\,\Gamma$ be designated $d^1, d^2, \ldots d^5$. From equations (6.11) and the designation of $A\,\Gamma$, and utilizing the form of H^*, Φ^*,

$$y_k = \begin{bmatrix} z_k^1 + d^1 w_k^* + d^3 w_{k+1}^* \\ z_k^2 + d^2 w_k^* \\ z_{k+1}^2 + d^3 w_k^* \\ z_{k+1}^2 \\ z_{k+2}^1 \end{bmatrix}$$

Define

$$P_k = \begin{bmatrix} z_k^1 \\ z_k^2 \\ z_{k+1}^1 \\ z_{k+1}^2 \\ z_{k+2}^1 \end{bmatrix} \qquad r_k^* = -\begin{bmatrix} d^1 w_k^* + d^3 w_{k+1}^* \\ d^2 w_k^* \\ d^3 w_k^* \\ 0 \\ 0 \end{bmatrix} \tag{8.5}$$

Thus

$$\begin{aligned} y_{k+1} &= \Phi^* y_k + A\,\Gamma\, w_k \\ P_k &= y_k + r_k^* \end{aligned} \tag{8.6}$$

From the definition of r_k^* it is seen that the vector will have correlation products of the form

$$E(r_k^* \, r_{k+j}^{*T}) \neq 0 \qquad \text{for } j = 1, 2, \ldots n \tag{8.7}$$

The spectral density of r_k^* is clearly a function of $A\,\Gamma$. Comparing (8.6) and (8.1)

$$\begin{aligned} y_{k+1} &= \Phi^* y_k + A\,\Gamma w_k^* \\ P_k &= y_k + r_k^* \end{aligned} \tag{8.6}$$

$$\begin{aligned} x_{k+1} &= \Phi x_k + w_k \\ z_k &= x_k + r_k \end{aligned} \tag{8.1}$$

it is noted that the format of the two sets of equations is identical. Thus, the net effect of forming "pseudo" state vectors from the sequential observations of dimensionality $\ell < n$ is seen, from (8.6), to be equivalent to adding correlated (colored) measurement noise r_k^* to the state vector y_k of an equivalent linear system, Φ^*.

The correlation products of r_k^* produce an apparent impasse in the application of the identification methods herein presented. The properties of $\hat{\phi}_1$ as developed in section 7 are dependent upon the assumption that the perturbing sequence have the property of statistical independence. The unbiased and minimum variance properties of $\hat{\phi}_1$, then, do not necessarily hold in the case of constrained observability. Computational investigation has verified that, in general, $\hat{\Phi}^*$ will not be an unbiased estimator.

Arriving at this same juncture, but through a different approach, Lee [3] suggested using only every n^{th} observation in forming the estimator $\hat{\Phi}^*$. However, from (8.5), it is apparent that no amount of time separation in the observations utilized will render r_k^* an uncorrelated sequence. Computational experimentation has verified this assertion, and is demonstrated in section 10.

Another approach considered was to extend the order of the estimator to include a coloring filter. Since the measurement noise is

lumped with the excitation to form ξ_1 in (8.3), and since there must exist a linear filter or order (n-1) that will produce r_k^* from a white spectrum, it was reasoned that an estimator $\hat{\Phi}^{**}$ of order 2n-1 might have the necessary combined characteristics of the system being identified and the coloring filter. This line of reasoning also appears to be invalid, however, as is demonstrated in section 10.

9. A COMPARISON WITH OTHER ESTIMATION SCHEMES

Ho and Walen [2] have presented the recursion formula

$$\hat{\Phi}_{m+1} = \hat{\Phi}_m + (1/\lambda\, m)(x_{m+1} - \hat{\Phi}_m x_m)\, x_m^T \tag{9.1}$$

where λ is a scalar constant. Using this formulation, the convergence

$$\lim_{m \to \infty} \hat{\Phi}_m = \Phi$$

can be demonstrated with probability 1, provided an appropriate value of λ is selected.

Compare this formulation with (4.5)

$$\hat{\Phi}_{m+1} = \hat{\Phi}_m + \alpha_m (x_{m+1} - \hat{\Phi}_m x_m)\, x_m^T P_m \tag{4.5}$$

$$\alpha_m = (x_m P_m x_m^T + 1)^{-1}$$

In (9.1) one can consider the term $(x_{m+1} - \hat{\Phi}_m x_m)\, x_m^T$ as being the current estimate (based on the transition from x_m to x_{m+1}) of the error in the estimator, $\hat{\Phi}_m$. Note that this term also appears in (4.5). Extending this line of reasoning, the factor $(1/\lambda m)$ in (9.1) is the weighting given this error term in forming the new estimator, $\hat{\Phi}_{m+1}$. Setting $\lambda = 1$ will cause every error so generated to be weighted equally in the aggregate estimate. For $\lambda < 1$, later terms will be weighted more heavily than the earlier terms.

The analogous weighting factor in (4.5) is the term $\alpha_m P_m$. The two formulations are seen to be identical except for this factor. Thus, (9.1) may be considered to be a special case of (4.5), with non-optimum weighting of the error term (optimality here being taken in the usual least squares sense). A comparison of the schemes with $\lambda = 1$ is presented in section 10.

Lee [3] considers the identification of a system with a scalar observable. The input distribution matrix and state transition matrix are constrained to be of the form

$$\Gamma = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ a \end{bmatrix} \qquad \Phi = \left[\begin{array}{c|c} 0 & \\ \cdot & I \\ \cdot & \\ \hline \multicolumn{2}{c}{\phi_n^T} \end{array}\right]$$

This system is described in section 6. Lee's formulation for the identification of this system is equivalent to that herein presented.

10. COMPUTATIONAL RESULTS

The purpose of this section is to obtain computational substantiation of the formulations herein presented. To this end, the following experimental objectives have been set forth:

1. Test the algorithm $\hat{\Phi} = R_1 R_0^{-1}$ for convergence to Φ.
2. Test the recursive algorithm for convergence.
3. Compare 1. and 2. above.
4. Test the recursive algorithm with additive measurement noise.
5. Investigate convergence under varying observability constraints.
6. Compare Ho's [2] method (equation 9.1) with the recursive algorithm.

Two model plants were selected to implement the testing. The first is a simple oscillator; the second a fourth order plant with two oscillatory modes, taken to be representative of the longitudinal dynamics of a large jet transport aircraft during normal cruise. [1]

Parameters for the two system are as follows:

2^{nd} order plant: $\omega = 1$ rad/sec $\quad T = \pi/8$ sec

4^{th} order plant: $\omega_1 = 1.15$ rad/sec $\quad \zeta_1 = .35$

$\omega_2 = .11$ rad/sec $\quad \zeta_2 = .035$

$T = 1.5$ sec

<u>Testing for convergence</u>. (tests 1, 2 and 3)

All the experimentation done supported the contention that $\hat{\Phi}$ does in fact converge to Φ, and that the estimator $\hat{\Phi}$ is therefore unbiased. The results of testing with the 2^{nd} order oscillator are presented as being typical. In fig. (3), the magnitude of the elements of $\hat{\Phi}$ are plotted as a function of the number of observations used in forming the estimator. These calculations were made using the nonrecursive (batch processing) algorithm.

In fig. (4), the elements of $\hat{\Phi}$ are plotted as a function of the number of iterations, and were calculated using the recursive algorithm. The recursive algorithm was initialized by setting $P_1 = 10^6 \cdot I$. Since it was demonstrated in section 7 that $\text{cov}(\hat{\phi}_i) = \sigma_i^2 P_m$, the large initial value for P_1 demonstrates this uncertainty in the initial value of $\hat{\Phi}_1$, taken to be the identity matrix. The resultant estimator is seen to be very nearly equivalent to that of the batch processing technique, by a comparison of figs. (3) and (4). Using this initialization scheme for the example shown, all elements of the recursive estimator were to within 4 significant figures of the nonrecursive estimator at 300 iterations. Identical data was used in the generation of the two estimators.

It is noted that the recursive form of the algorithm offers several computational advantages, including its suitability for real time implementation and the fact that no matrix inversion is required in its implementation. Because of these advantages, the recursive form was used for the majority of the remaining testing, the author being satisfied that no degradation of the estimator $\hat{\Phi}$ would result from so doing.

$$\Phi = \begin{bmatrix} \cos\frac{\pi}{8} & \sin\frac{\pi}{8} \\ -\sin\frac{\pi}{8} & \cos\frac{\pi}{8} \end{bmatrix}$$

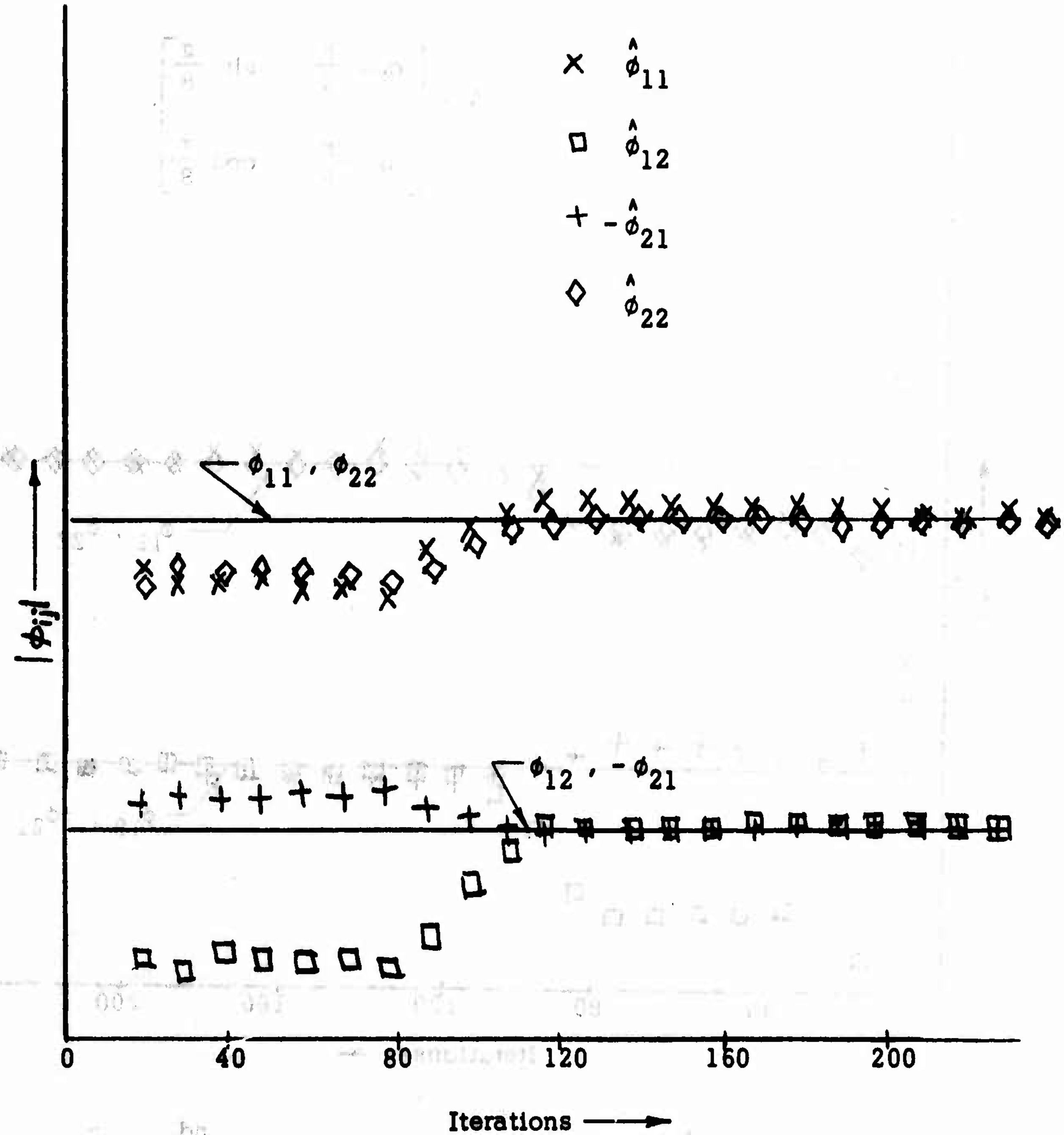


Figure 3. Estimator $\hat{\Phi}$, using the nonrecursive algorithm, 2nd order plant.

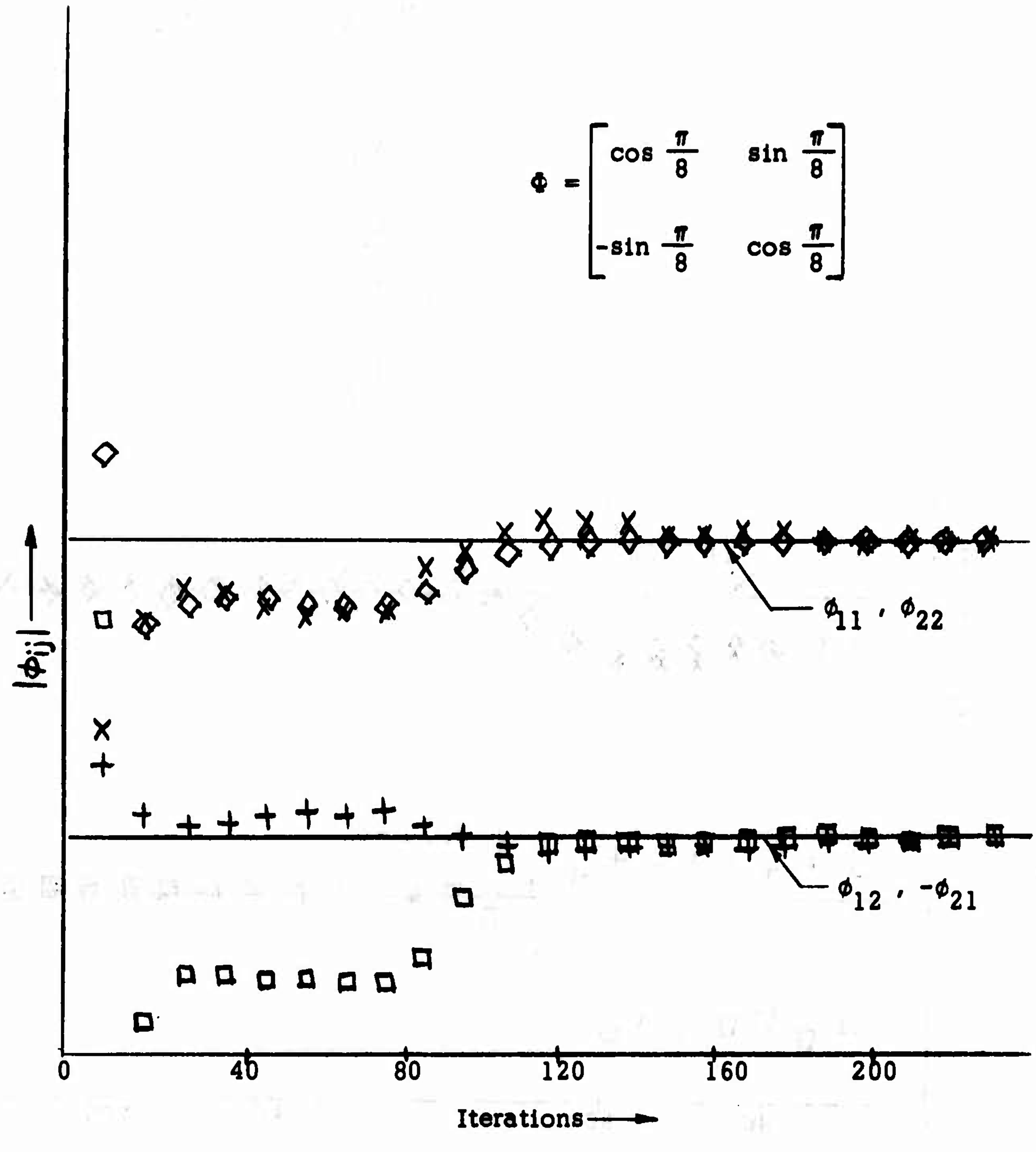


Figure 4. Estimator $\hat{\phi}$, using the recursive algorithm, 2nd order plant.

Testing with measurement noise. (test 4)

For the remaining tests, especially with the higher order system, it is desirable to introduce a scalar measure of merit for the estimators. While the definition of such a measure is certainly arbitrary, it is the author's opinion that such a measure should make a comparison of the characteristic roots of the estimator with the characteristic roots of the actual system. This method seems especially appropriate when comparing $\hat{\Phi}*$ matrices, the make up of which vary considerably with varying observability constraints. By definition, then,

$$M_1 = \sum_{i=1}^{n} \frac{|\lambda_i - \hat{\lambda}_i|^2}{|\lambda_i|^2}$$

where the λ_i are the characteristic roots of the matrix Φ, and the $\hat{\lambda}_i$ the roots of the estimator $\hat{\Phi}$.

A typical result of estimating with uncorrelated measurement noise, using the 4^{th} order model, is shown in Fig. (5). Here M_1 for the case of noisy observations is compared with an estimator formed from the same data, but without the additive measurement noise. For this example, the ratio of variances for excitation and measurement noise was taken to be

$$4 < \sigma^2_{w_i} / \sigma^2_{r_j} < 10 \qquad \text{for } i, j = 1, 2, 3, 4.$$

Testing with constrained observability. (test 5.)

As stated in section 8, the estimators $\hat{\Phi}*$, formed from systems having constrained observability, in general are not unbiased. However, several interesting properties of the estimator have come into evidence from computational experimentation.

The first experiment under constrained observability was conducted with the 2^{nd} order oscillator, and the estimator $\hat{\Phi}*$ formed from observations on a single mode of the system.

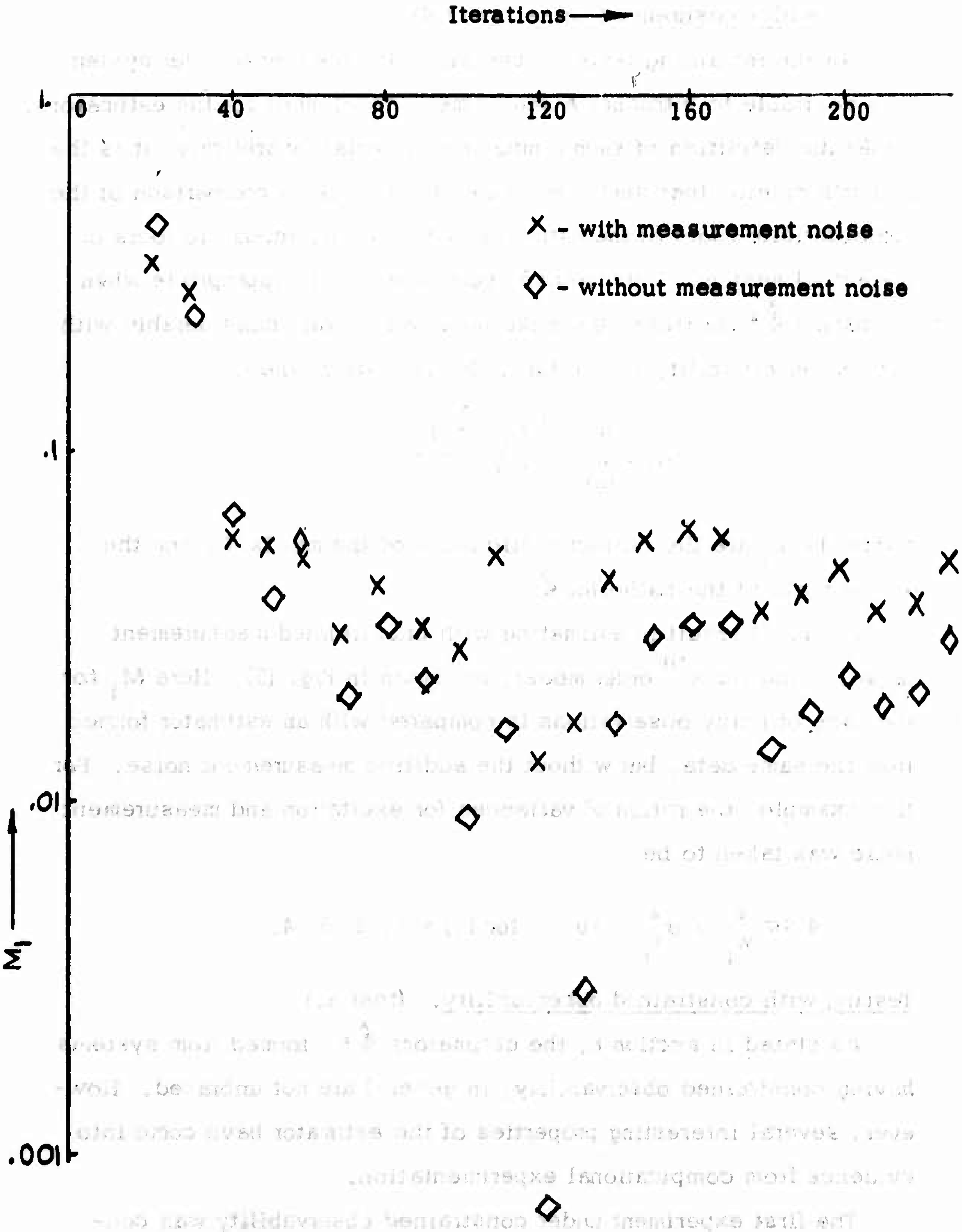


Figure 5. Comparison of estimators formed with and without measurement noise, 4th order plant.

$$\sigma_w^2/\sigma_1^2 \sim 10. \quad M_1 = \Sigma \frac{|\lambda_i - \hat{\lambda}_i|^2}{|\lambda_i|^2}$$

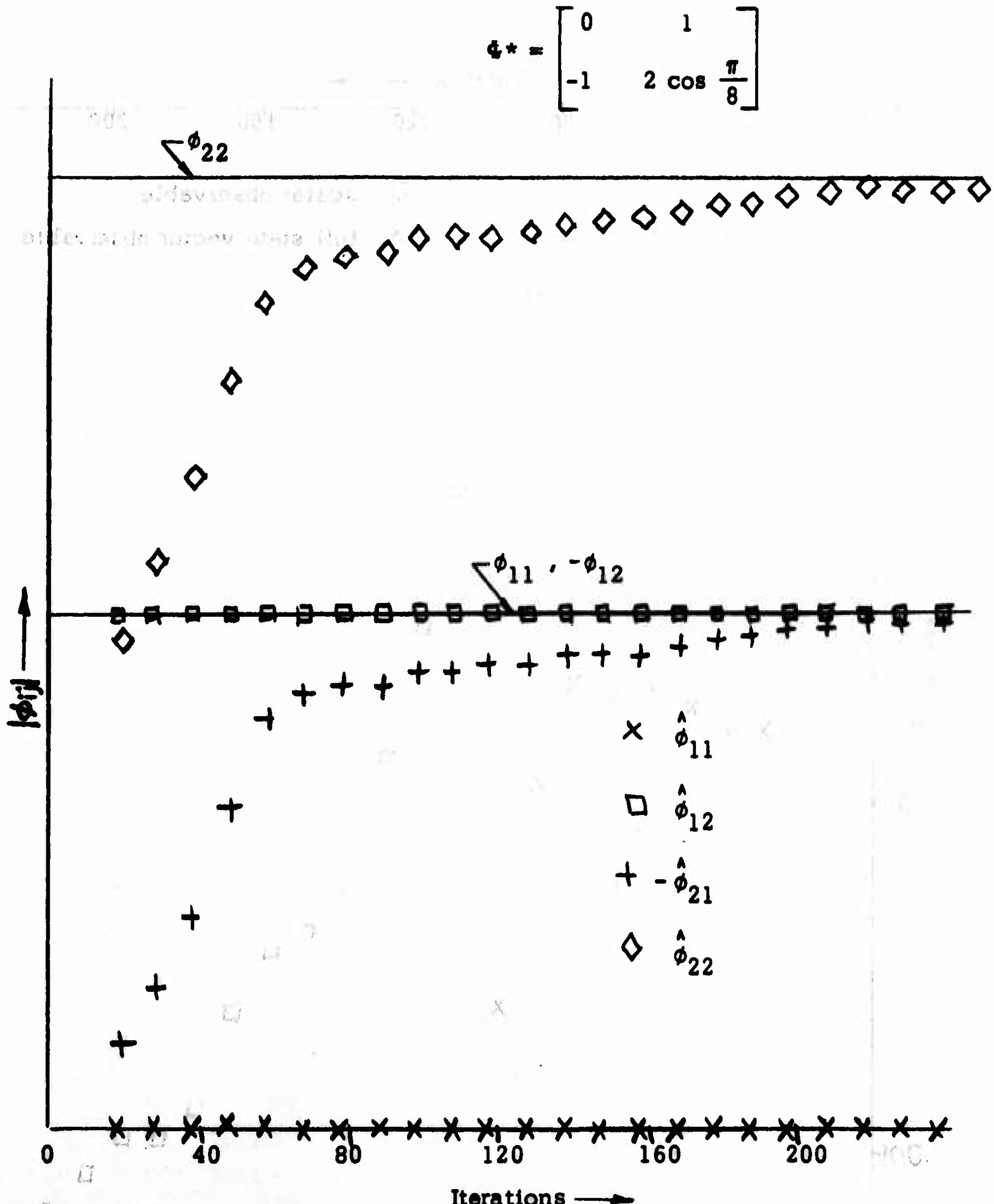


Figure 6. Estimator $\hat{\phi}^*$ for a simple oscillator. Estimator formed from observations on a single mode of the system.

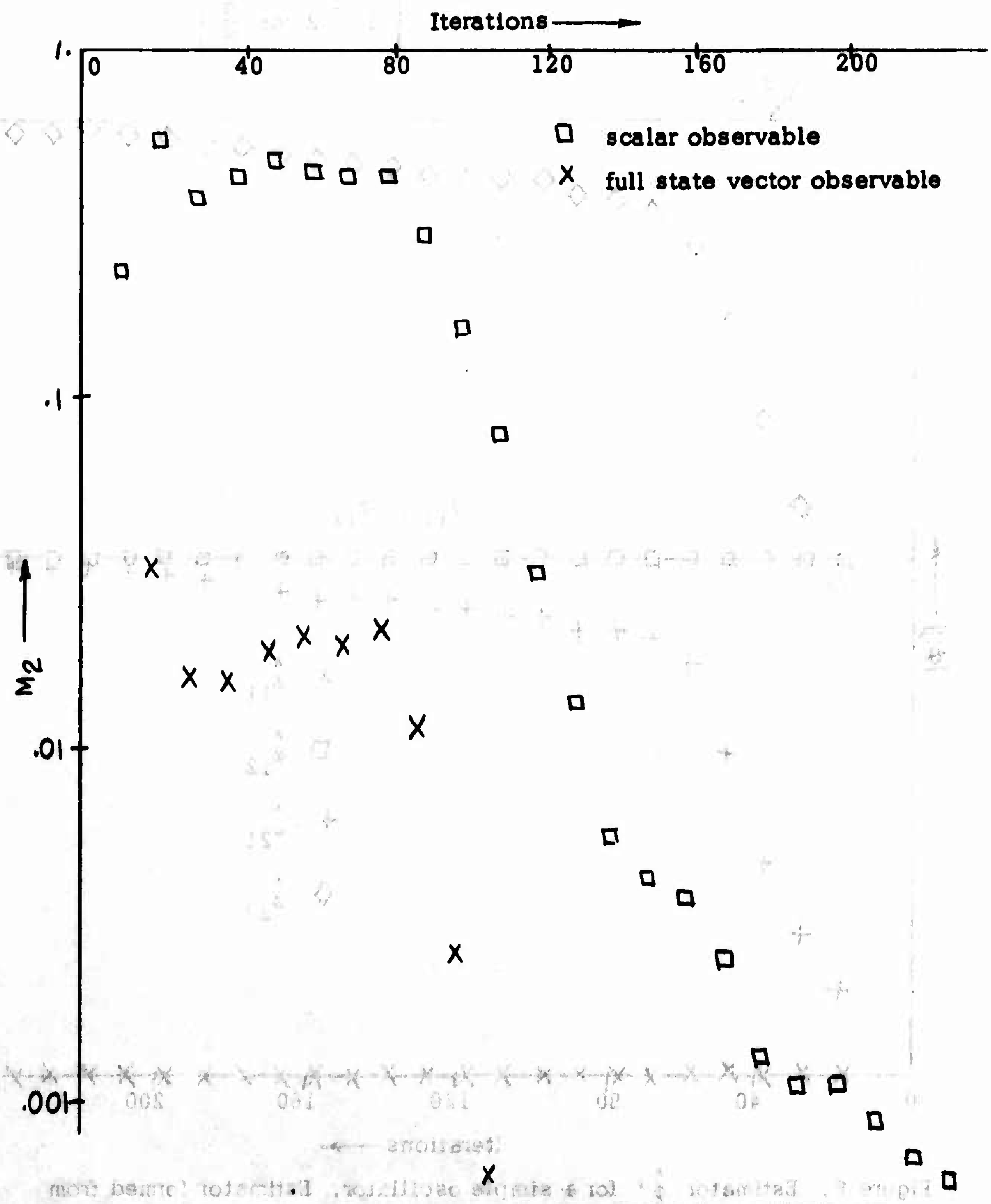


Figure 7. Comparison of convergence for estimators formed from constrained and unconstrained observations, 2nd order system.

$$M_2 = \frac{\| \Phi - \hat{\Phi} \|}{\| \Phi \|}$$

Results of the experiment are shown in figs. (6) and (7). Convergence to $\Phi *$, or to a matrix very close to $\Phi *$, is demonstrated in this case. Fig. (7) shows a comparison of the constrained observability estimator, $\hat{\Phi} *$, and the estimator formed from the full state vector, $\hat{\Phi}$. The measure of merit suggested by Lee [3]

$$M_2 = \frac{\| \Phi - \hat{\Phi} \|^2}{\| \Phi \|^2}$$

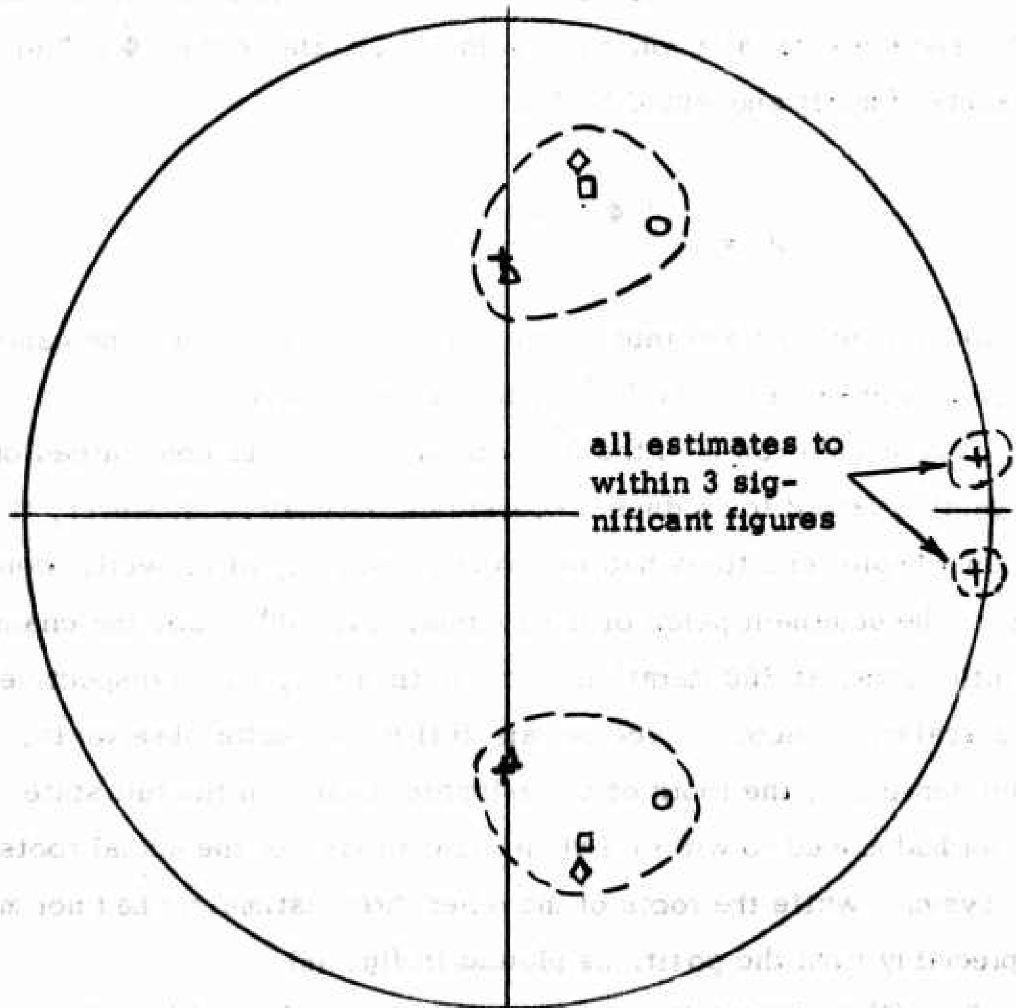
is plotted here. At 300 interations, M_2 equalled 10^{-6} for the case of the vector observable; 5×10^{-4} for the scalar observable.

Experimentation with the 4^{th} order system under constrained observability failed to produce a convergent estimator. However, five distinct implementations had the common property of correctly identifying the dominant poles of the system. Fig. (8) shows the characteristic roots, at 300 iterations, of 4 estimators, based respectively on a scalar, 2 vector, 3 vector, and full state vector observable. At 2000 iterations, the roots of the estimator based on the full state vector had moved to within 3 significant figures of the actual roots of the system, while the roots of the other three estimators had not moved appreciably from the positions plotted in fig. (8).

Fig. (9) compares the roots of an estimator formed by using only every 4^{th} observation with an estimator formed by a scalar observable at each sampling instant. Using only every 4^{th} sample is the method suggested by Lee [3], as discussed in section 8. Note here that the state transition for 4 sampling instants is given by $\Phi_2^*(4T) = [\Phi_2^*(T)]^4$. The roots of $\Phi_2^*(4T)$ are then $(\lambda_i)^4$, where λ_i are the roots of $\Phi_2^*(T)$. Making this calculation places the dominant poles of $\hat{\Phi}_2^*(T)$ in close proximity of the poles of Φ, as shown in fig. (10).

In section 8 it was suggested that the order of the estimator $\hat{\Phi}**$ for the case of constrained observability be increased to 2n-1 to allow for the inclusion of a coloring filter. A typical result of such an augmented estimator is shown in fig. (11). Also shown, for comparison,

Z PLANE PLOT OF CHARACTERISTIC ROOTS



+ Roots of the actual state transition matrix

○ Roots of $\hat{\Phi}^*_1$, based on a scalar observable

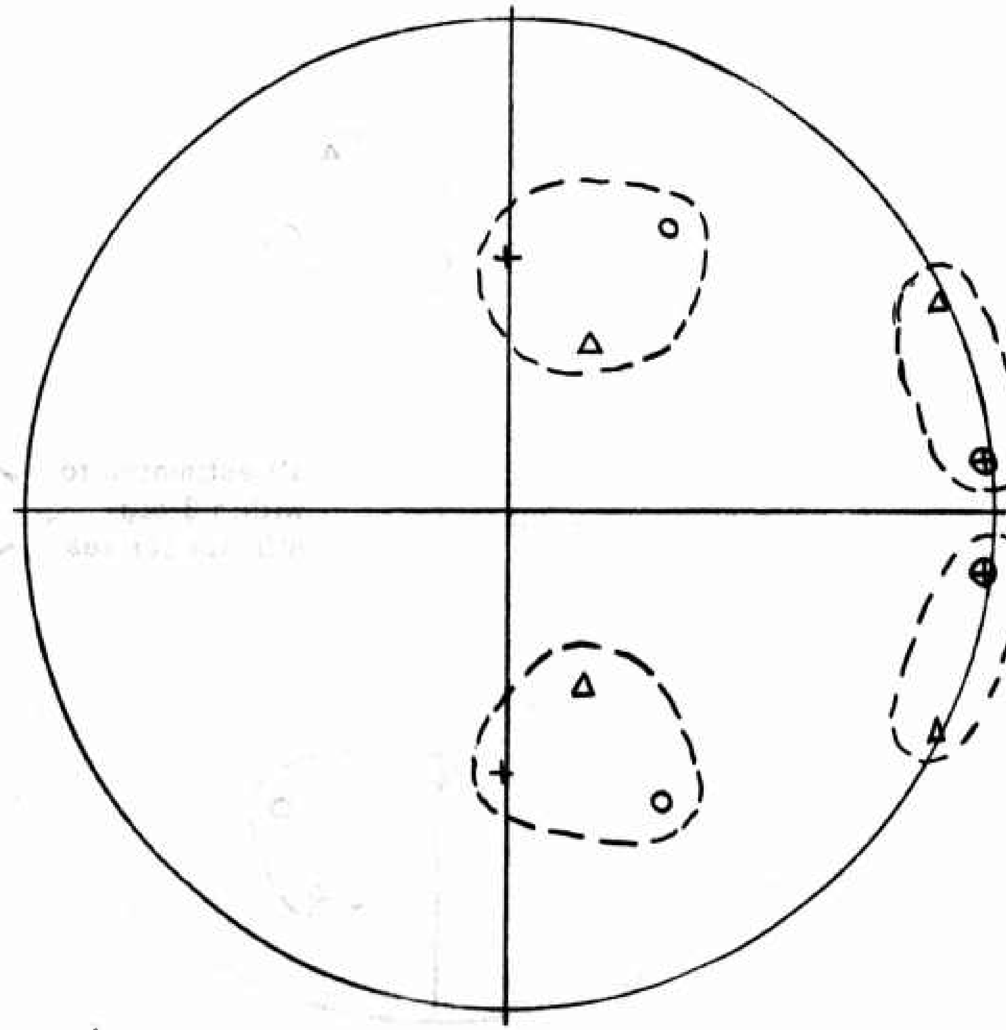
◇ Roots of $\hat{\Phi}^*_2$, based on a 2 vector observable

□ Roots of $\hat{\Phi}^*_3$, based on a 3 vector observable

Δ Roots of $\hat{\Phi}$. All state modes observable

Figure 8. The characteristic roots of 4 estimators are compared. Only the estimator $\hat{\Phi}$ is converging to Φ. The dominant (low frequency) roots of all estimators are within 3 significant figures of the dominant roots of Φ. 300 iterations were used to generate all estimators.

Z PLANE PLOT OF CHARACTERISTIC ROOTS



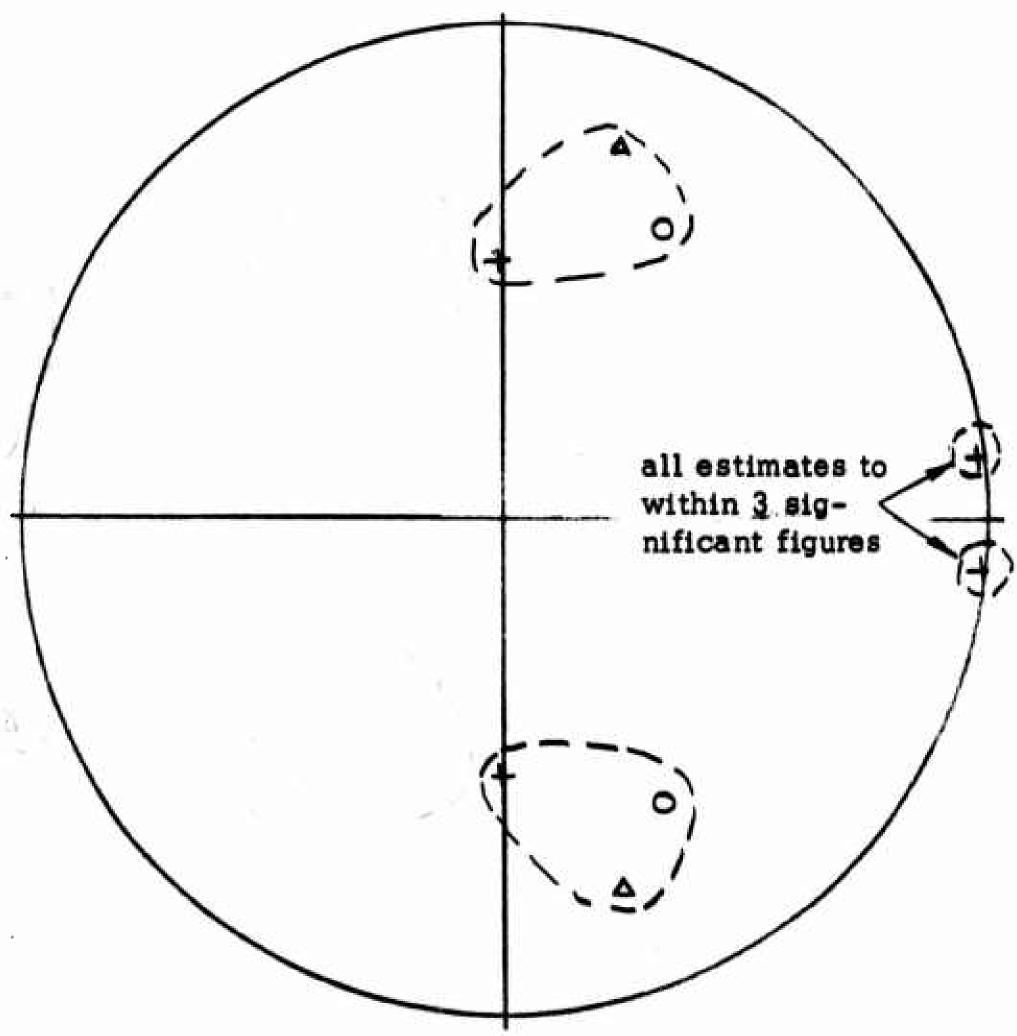
+ Roots of the actual state transition matrix

o Roots of $\hat{\Phi}_1^*$, based on a scalar observable

Δ Roots of $\hat{\Phi}_2^*$, based on a scalar observable, using only every 4^{th} observation

Figure 9. The characteristic roots of an estimator $\hat{\Phi}_2^*$, formed from every 4^{th} observation, are compared with the roots of $\hat{\Phi}_1^*$, formed from successive scalar observations. 2000 iterations were used to generate the estimators. The dominant (low frequency) roots of $\hat{\Phi}_1^*$ are within 3 significant figures of the roots of Φ .

Z PLANE PLOT OF CHARACTERISTIC ROOTS



+ Roots of the actual state transition matrix

O Roots of $\hat{\Phi}_1^*$

Δ Roots of $\hat{\Phi}_2^*(T)$

Figure 10. The characteristic roots of $\hat{\Phi}_2^*(T)$ are generated from $\hat{\Phi}_2^*(4T)$. These roots are compared with the roots of $\hat{\Phi}_1^*$. The data of fig. (9) was used to calculate these roots.

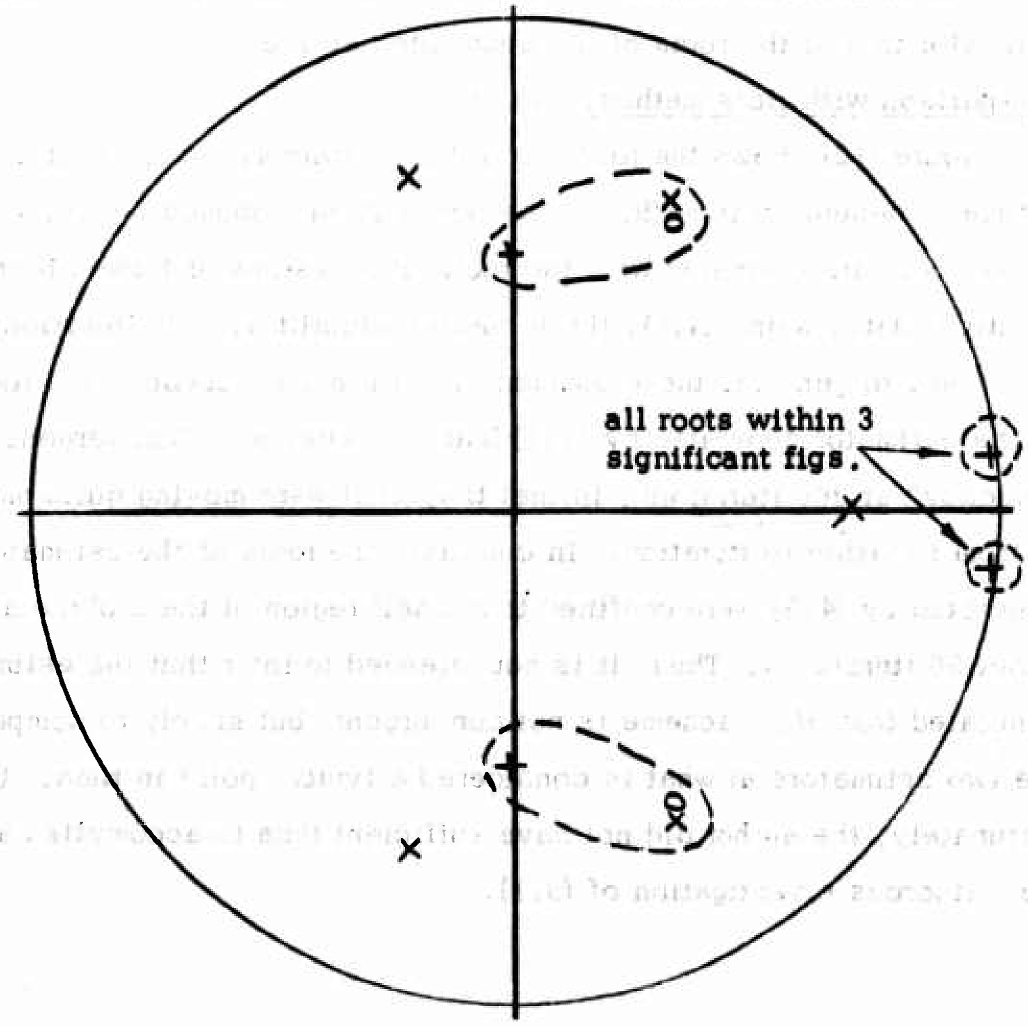
are the roots of a 4^{th} order estimator based on a scalar observable. 2000 iterations were used to generate the data shown.

It is of interest to note the proximity of the roots of the 4^{th} order estimator to 4 of the roots of the augmented estimator.

Comparison with Ho's method. (test 6.)

Figure (12) shows the location of the characteristic roots of an estimator generated from (9.1), the formulation proposed by Ho. [2] These roots are compared with the roots of an estimator formed from identical data, using (4.5), the recursive algorithm. 300 iterations were used to generate these estimators. In this simulation, the roots of the estimator generated by (9.1) had not acquired a "convergent tendency" at 300 iterations, in that they still were moving quite markedly from iteration to iteration. In contrast, the roots of the estimator generated by (4.5) were confined to a small region of the z plane after about 50 iterations. Thus, it is not intended to imply that the estimator generated from Ho's scheme is not convergent, but simply to compare the two estimators at what is considered a typical point in time. Unfortunately, the author did not have sufficient time to accomplish a more rigorous investigation of (9.1).

Z PLANE PLOT OF CHARACTERISTIC ROOTS



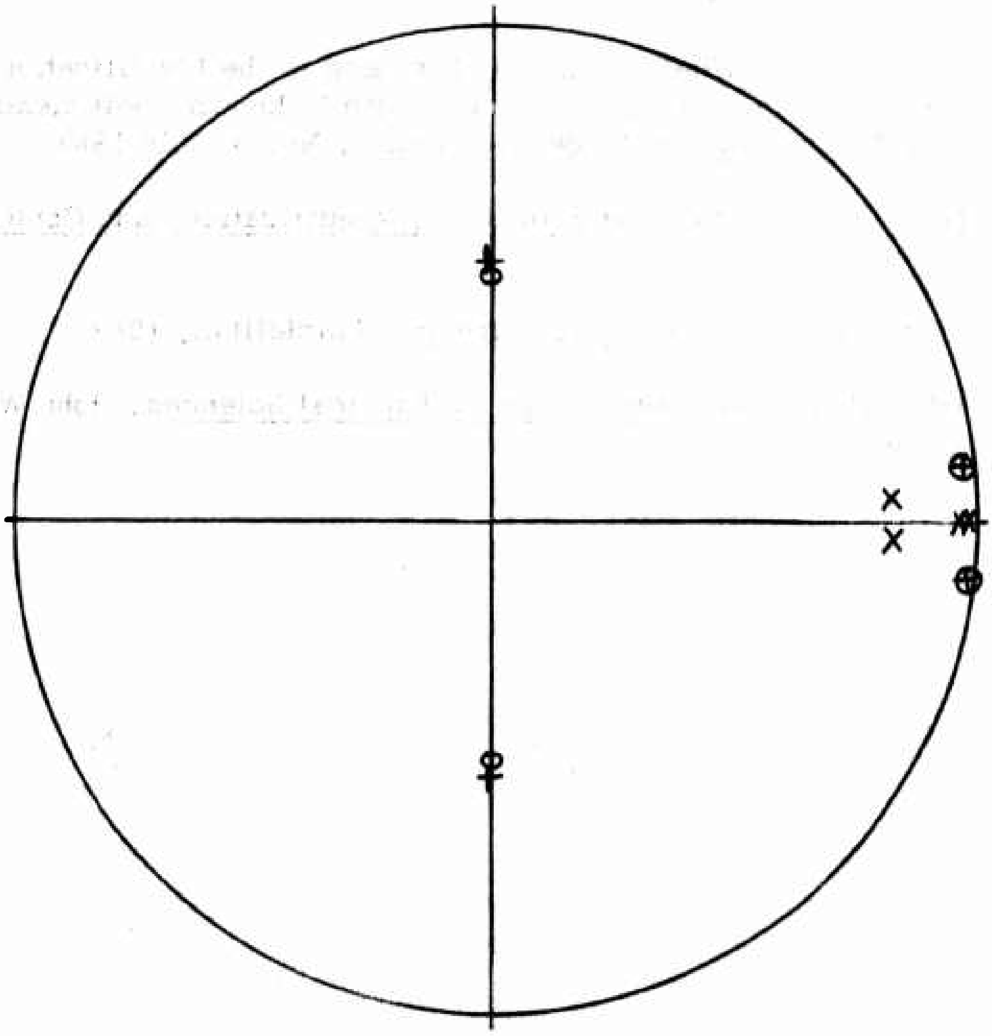
+ Roots of the actual state transition matrix

o Roots of $\hat{\Phi}_1^*$, based on a scalar observable

× Roots of $\hat{\Phi}^{**}$, augmented to order 7

Figure 11. The characteristic roots of an augmented estimator $\hat{\Phi}^{**}$, formed from a scalar observable, are compared with the roots of $\hat{\Phi}_1^*$. 2000 iterations were used to generate the estimators.

Z PLANE PLOT OF CHARACTERISTIC ROOTS



+ Roots of the actual state transition matrix

o Roots of $\hat{\Phi}$. Full state vector observable

× Roots of Ho's estimator

Figure 12. The characteristic roots of an estimator proposed by Ho are compared with the estimator $\hat{\Phi}$, formed from the full state vector. 300 iterations were used to generate the estimators.

BIBLIOGRAPHY

1. Blakelock, J. H. Automatic Control of Aircraft and Missiles. John Wiley & Sons, 1962

2. Ho, Y.C. and Whalen, B., "An Approach to the Identification and Control of Linear Dynamic Systems with Unknown Coefficients", IEEE Transactions on Automatic Control, No. 3, July 1963

3. Lee, R. C. K. Optimal Estimation, Identification, and Control. MIT Press, 1964

4. Lindgren, B. W. Statistical Theory. MacMillan, 1962

5. Wilf, H. S. Mathematics for the Physical Sciences. John Wiley & Sons, 1962

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE

Identification of Linear Sampled Data Systems

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Masters thesis - June 1967

5. AUTHOR(S) (Last name, first name, initial)

Blackner, Ronald K. Lieutenant Commander U. S. Navy

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1967	46	5

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. AVAILABILITY/LIMITATION NOTICES This document is subject to special export controls and each transmittal to foreign government or foreign nationals may be made only with prior approval of the U. S. Naval Postgraduate School.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Ordnance Test Station, China Lake, California P.O. #6-0048

13. ABSTRACT

A least squares estimator $\hat{\Phi}$ is derived for the state transition matrix Φ of a linear, stationary sampled data system operating in a stochastic environment. The estimator $\hat{\Phi}$ is shown to be unbiased and minimum variance under the condition of full observability of the state vector of the system. The estimator is also shown to be the Maximum Likelihood Estimator for the case of the stochastic environment having Gaussian statistics. The estimation scheme is compared with two other recently published estimation schemes, both of which are shown to be special cases of the scheme herein presented.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Estimation						
Identification						
Stochastic						
Discrete linear models						